Research Policy 45 (2016) 593–603

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Evaluating research: A multidisciplinary approach to assessing research practice and quality

Pär Mårtensson[a,1], Uno Fors[b,2], Sven-Bertil Wallin[c,3], Udo Zander[d,4], Gunnar H Nilsson[e,*]

[a] Department of Management and Organization, Stockholm School of Economics, Sveavägen 65, SE-113 83 Stockholm, Sweden
[b] Department of Computer and Systems Sciences, Stockholm University, PO Box 7003, SE-164 07 Kista, Sweden
[c] Almain AB, Lilla Norrby, SE-178 91 Munsö, Sweden
[d] Department of Marketing and Strategy, Stockholm School of Economics, Sveavägen 65, SE-113 83 Stockholm, Sweden
[e] Department of Neurobiology, Care Sciences and Society, Karolinska Institutet, SE-171 77 Stockholm, Sweden

ARTICLE INFO

Article history:
Received 17 June 2014
Received in revised form
23 November 2015
Accepted 24 November 2015

Keywords:
Research practice
Research quality
Research evaluation
Multidisciplinary
Concept model
Rigour
Relevance

ABSTRACT

There are few widely acknowledged quality standards for research practice, and few definitions of what constitutes good research. The overall aim was therefore to describe what constitutes research, and then to use this description to develop a model of research practice and to define concepts related to its quality. The primary objective was to explore such a model and to create a multidisciplinary understanding of the generic dimensions of the quality of research practice. Eight concept modelling working seminars were conducted. A graphic representation of concepts and their relationships was developed to bridge the gap between different disciplines. A concept model of research as a phenomenon was created, which included a total of 18 defined concepts and their relationships. In a second phase four main areas were distilled, describing research practice in a multidisciplinary context: *Credible*, *Contributory*, *Communicable*, and *Conforming*. Each of these was further specified in a concept hierarchy together with a defined terminology. A comprehensive quality model including 32 concepts, based on the four main areas, was developed for describing quality issues of research practice, where the model of research as a phenomenon was used to define the quality concepts. The quality model may be used for further development of elements, weights and operationalizations related to the quality of research practice in different academic fields.

1. Introduction

1.1. Background

A fundamental question that can be posed within any field of research is 'What constitutes good or high quality *research* (*or scientific*) *practice*?'. This question is relevant for research both in a university context and in an organizational or innovation context for research and development activities. However, before such criteria can be formulated, we need a reasonably common understanding of what research itself really is.

Science and research are ontologically challenging, and previous research reveals different views and remains ambiguous. A recent definition of science was proposed by the British Science Council: "Science is the pursuit of knowledge and understanding of the natural and social world following a systematic methodology based on evidence" (Science Council, 2009: www.sciencecouncil.org/definition). Based on somewhat similar definitions, several studies have explored the concept of *research*. In this respect Israel (2005) acknowledged and explored the complexity of science, Patton (1990) mentioned that it is important to identify the purpose of research, and Gall et al. (1996) discussed how research might contribute in the field of education. In the medical domain, Grinnell (1990) argued that the endings of clinical research protocols are of importance in distinguishing therapy from research. In 2000 the same author stated that the everyday practice of science is neither realism nor social constructivism, but rather is balanced on a contextual ledge between the two, and said that he considered discovery and credibility to be the two central features of research (Grinnell, 2000). Ulrich (2006) has analysed different traps that are currently common and that lead to a somewhat limited reflective

* Corresponding author. Tel.: +46 8 524 800 00.
E-mail addresses: par.martensson@hhs.se (P. Mårtensson), uno@dsv.su.se (U. Fors), sven-bertil.wallin@almain.se (S.-B. Wallin), udo.zander@hhs.se (U. Zander), gunnar.nilsson@ki.se (G.H. Nilsson).
[1] Tel.: +46 8 736 90 00.
[2] Tel.: +46 8 674 7479.
[3] Tel.: +46 70 219 91 90.
[4] Tel.: +46 8 736 90 00.

http://dx.doi.org/10.1016/j.respol.2015.11.009

research practice, and he describes a rethinking approach. Further, Quaye (2007) argues for extending what counts as research within the social sciences so that it is more likely to include different methodologies and writing genres. Nickelsen (2009), in a similar approach, supports the notion of interventionist research that is not just focused on simple one-way causation in the field that is being studied. In parallel with this, there has been ongoing discussion about rethinking knowledge production in general (e.g. Hessels and van Lente, 2008; Tsao et al., 2008). In this new mode of knowledge production, often referred to as Mode 2, knowledge is produced in the context of an application (Gibbons et al., 1994). Knowledge can be produced in different contexts, and the concepts of 'knowing in action' (Amin and Roberts, 2008) and 'situated learning' (Lave and Wenger, 1991) highlight the importance of a variety of contextual factors. It is important to keep this in mind, not least in the light of the considerable amount of knowledge production taking place in Research and Development (R&D) departments in companies.

In summary, there is broad criticism of the so-called linear model of science, and it is argued that concepts such as intuition and passion have become just as important as objectivity and logic (Dash, 2009; Grinnell, 2009), and there are several very different views on, and definitions of, research practice. In our paper we therefore concentrate our efforts on working towards a generic definition (or model) of what research is. Then, based on this model, it may be possible to define the generic components of quality of research practice.

However, before describing and discussing our study and the resulting model, we need to clarify the terminology used. In some publications and websites on this topic, there seems to be some confusion between the term *research* and the term *science*, and these terms seem to be used interchangeably. In our view the term science is broader, and research is more like the *practice* of working in a scientific manner. Research is what you practise, and the result of this work is science. We have used the terms *research* and *research practice* throughout this paper, as the scope of our study comprises trying to define what high quality science production might be.

Moreover, as the evaluation of research practice is one of our end-goals, we may also need to define what we mean by *evaluation*. In our view, the practice of evaluation can be defined as an activity in which certain aspects of the quality of research practice are investigated. But what does this really mean? The ambition to evaluate research has a long history that is full of tensions, ambiguities and misunderstandings. Some countries have formed national commissions for evaluating research, which seem to focus on bibliometric analyses to measure research quality (Jiménez-Contreras et al., 2003), but the evaluation of research may include many other aspects. The current debate is for example highlighting the problem of having evaluations "led by the data rather than by judgement" (Hicks et al., 2015: 429). An often cited definition of evaluation is "...a process for collecting and synthesizing evidence that can make conclusions about the state of affairs, value, merit, worth, significance or quality of a program, product ..." (Mathison, 2005), which implies that evaluation can use numerous methods and measure a wide variety of aspects (see also Mertens, 2015).

1.2. Dimensions of the quality of research practice

Evaluation of the quality of research practice is a truly important issue in most scientific domains and at many levels (European Science Foundation, 2012). Increasingly, we are also seeing these assessment efforts across disciplinary and national boundaries. More or less elaborate efforts have been made in recent years to evaluate the quality of research practice in a host of different settings. These efforts affect resource allocation, scientific activity, and the very lives of researchers across the globe. Quality is the focus for several different reasons, and is examined in a variety of contexts such as in the evaluation of:

- research grant applications
- research manuscripts and publications
- specific research topics
- research groups and constellations
- institutions
- national systems for producing science and innovation

Regarding the issue of measuring the quality of research in the wider scientific community, it is difficult to find a universal definition of what constitutes good scientific practice. The focus at some universities is only on the number and quality of publications in scientific journals, whereas other institutions focus on all kinds of publications. However, in an increasing number of academic fields it is becoming more and more common for scientific output to be measured in ways other than simply counting the number and quality of publications.

Several costly quality-assessment projects have been undertaken lately to improve the quality of research practice at the authors' own institutions, to determine which research areas should receive funding, to find out whether and where quality improvements are necessary, and to benchmark the quality of a certain institution against that of leading international institutions. However, the available scientific literature on research quality, and on what can really be defined as research, is scarce. Some examples do exist. In Italy, for example, national reference guidelines for the evaluation of research practice have, in general, advocated an approach that includes socio-economic impact, resource attraction and resource management as criteria (CIVR, 2006). In the US, the criteria for evaluating research grant applications at the National Institutes of Health include short definitions of five concepts: significance, approach, innovation, investigators, and environment (NIH, 2008). In a recent evaluation of research constellations within a large university in Sweden, the quality of research practice was measured by considering the attention received concerning the scientific, technological, clinical and socio-economic significance of their publications, including the implementation of research results in society (External Research Assessment (ERA), 2010). In Sweden today, however, there seem to be at least as many ways to measure what constitutes a good scientific study or publication, as there are research institutions. In Canada, standard quality assessment criteria for research papers have been developed, and these deal separately with quantitative and qualitative research studies (Kmet et al., 2004).

However, it is not our goal to distinguish some types of scientific methods that are inherently 'good' from others that may be 'bad'. Our contention is that almost any scientific method can be appropriate, given a sound research design. It is the research question(s) at hand that should lead to the decision on which research design and method(s) should be used, and quality may be high as long as the methods are used with rigour and quality. In our view, theories can be seen as 'maps' and research methods as 'nets'; both are highly context dependent in how they find and capture the elements for producing new knowledge.

Quality assurance and evaluation measures are meant to be as objective and reliable as possible. They generally have the aim of increasing awareness about the current status and standing of the research that is underway. However, the general problem is that nearly all the recent evaluation projects have used different measures and weights for the applied variables, making it difficult to compare an institution's evaluation results with those of other institutions or disciplines. Specific examples of proxy variables that